# RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: _10/560,760_

Source: _1 FWP_

Date Processed by STIC: _2/22/07_

# *ENTERED*

IFWP

RAW SEQUENCE LISTING                    DATE: 02/22/2007
PATENT APPLICATION:  US/10/560,760      TIME: 12:11:02

Input Set : A:\A04002\G3111EP.ST25.txt
Output Set: N:\CRF4\02222007\J560760.raw

```
 3 <110> APPLICANT: Evologic S.A.
 4        Marliere, Phillipe
 6 <120> TITLE OF INVENTION: Cloning of gluconate dehydratase gcnD gene
 8 <130> FILE REFERENCE: G 3111 EP
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/560,760
C--> 10 <141> CURRENT FILING DATE: 2005-12-14
 10 <160> NUMBER OF SEQ ID NOS: 17
 12 <170> SOFTWARE: PatentIn version 3.1
 14 <210> SEQ ID NO: 1
 15 <211> LENGTH: 1812
 16 <212> TYPE: DNA
 17 <213> ORGANISM: Agrobacterium tumefaciens
 19 <220> FEATURE:
 20 <221> NAME/KEY: CDS
 21 <222> LOCATION: (1)..(1809)
 22 <223> OTHER INFORMATION:
W--> 25 <400> 1
 26 atg acg aca tct gat aat ctt cct gca act cag ggc aag ctc cgt tcg      48
 27 Met Thr Thr Ser Asp Asn Leu Pro Ala Thr Gln Gly Lys Leu Arg Ser
 28 1               5                   10                  15
 30 cgc gcc tgg ttc gac aac cca gcc aat gcg gac atg acc gcg ctt tat      96
 31 Arg Ala Trp Phe Asp Asn Pro Ala Asn Ala Asp Met Thr Ala Leu Tyr
 32              20                  25                  30
 34 ctc gag cgt tac atg aac ttc ggt ctc agc cag gcc gag ctt cag tcc     144
 35 Leu Glu Arg Tyr Met Asn Phe Gly Leu Ser Gln Ala Glu Leu Gln Ser
 36          35                  40                  45
 38 gac cgc ccg att atc ggt att gcg cag acc ggt tcc gac ctt tcg ccc     192
 39 Asp Arg Pro Ile Ile Gly Ile Ala Gln Thr Gly Ser Asp Leu Ser Pro
 40      50                  55                  60
 42 tgc aac cgt cat cat ctg gaa ctc gcc aac cgt ctg cgt gaa ggc att     240
 43 Cys Asn Arg His His Leu Glu Leu Ala Asn Arg Leu Arg Glu Gly Ile
 44 65                  70                  75                  80
 46 cgt gaa gcc ggc ggc atc gcc atc gaa ttc ccg gtg cat ccg atc cag     288
 47 Arg Glu Ala Gly Gly Ile Ala Ile Glu Phe Pro Val His Pro Ile Gln
 48                  85                  90                  95
 50 gaa acc ggt aag cgt ccg aca gcg ggc ctt gat cgc aac ctg gct tac     336
 51 Glu Thr Gly Lys Arg Pro Thr Ala Gly Leu Asp Arg Asn Leu Ala Tyr
 52                  100                 105                 110
 54 ctc ggc ctc gtg gaa gtg ctt tat ggc tat ccg ctc gac ggc gtt gtt     384
 55 Leu Gly Leu Val Glu Val Leu Tyr Gly Tyr Pro Leu Asp Gly Val Val
 56              115                 120                 125
 58 ctg acc atc ggc tgc gac aag acc acg cct gcc tgt ctt atg gcg gcg     432
 59 Leu Thr Ile Gly Cys Asp Lys Thr Thr Pro Ala Cys Leu Met Ala Ala
```

```
60      130                 135                 140
62 gcc acc gtc aac att ccg gcc atc gcc ctt tcc gtc ggt ccc atg ctg      480
63 Ala Thr Val Asn Ile Pro Ala Ile Ala Leu Ser Val Gly Pro Met Leu
64 145                 150                 155                 160
66 aac ggc tgg ttc cgc ggt gag cgc acc ggt tcc ggc acc atc gtc tgg      528
67 Asn Gly Trp Phe Arg Gly Glu Arg Thr Gly Ser Gly Thr Ile Val Trp
68                 165                 170                 175
70 aag gcc cgc gaa ctg ctg gcg aag ggc gag atc gat tac cag ggc ttc      576
71 Lys Ala Arg Glu Leu Leu Ala Lys Gly Glu Ile Asp Tyr Gln Gly Phe
72                 180                 185                 190
74 gtc aag ctc gtt gcc tcg tct gcc ccg tcc acc ggc tat tgc aac acc      624
75 Val Lys Leu Val Ala Ser Ser Ala Pro Ser Thr Gly Tyr Cys Asn Thr
76          195                 200                 205
78 atg ggc acg gca aca acc atg aac tcg ctc gcc gaa gcg ctc ggc atg      672
79 Met Gly Thr Ala Thr Thr Met Asn Ser Leu Ala Glu Ala Leu Gly Met
80      210                 215                 220
82 cag ctt ccc ggc tcc gcc gcc att ccg gcg cct tac cgt gac cgt cag      720
83 Gln Leu Pro Gly Ser Ala Ala Ile Pro Ala Pro Tyr Arg Asp Arg Gln
84 225                 230                 235                 240
86 gaa gtc tct tac ctc acc ggc ctg cgc atc gtc gac atg gtc agg gaa      768
87 Glu Val Ser Tyr Leu Thr Gly Leu Arg Ile Val Asp Met Val Arg Glu
88                 245                 250                 255
90 gac ctg aaa cca tca gac atc atg acc aag gat gcc ttc atc aac gcc      816
91 Asp Leu Lys Pro Ser Asp Ile Met Thr Lys Asp Ala Phe Ile Asn Ala
92                 260                 265                 270
94 atc cgc gtt aat tcg gcg atc ggc ggt tcc acc aac gcg ccg atc cat      864
95 Ile Arg Val Asn Ser Ala Ile Gly Gly Ser Thr Asn Ala Pro Ile His
96          275                 280                 285
98 cta aac ggc ctt gcc cgc cat gtc ggc gtc gag ctg acg gtg gat gac      912
99 Leu Asn Gly Leu Ala Arg His Val Gly Val Glu Leu Thr Val Asp Asp
100     290                 295                 300
102 tgg cag acc tat ggc gaa gac gtg ccg ctg ctc gtc aac ctg cag ccg      960
103 Trp Gln Thr Tyr Gly Glu Asp Val Pro Leu Leu Val Asn Leu Gln Pro
104 305                 310                 315                 320
106 gca ggc gaa tat ctc ggc gag gac tat tac cat gcc ggc ggc gtt ccc     1008
107 Ala Gly Glu Tyr Leu Gly Glu Asp Tyr Tyr His Ala Gly Gly Val Pro
108                 325                 330                 335
110 gct gtc gtc aat cag ctg atg acc caa ggg ctg atc atg gaa gac gcc     1056
111 Ala Val Val Asn Gln Leu Met Thr Gln Gly Leu Ile Met Glu Asp Ala
112             340                 345                 350
114 atg acc gtc aac ggc aag acc atc ggc gac aat tgc cgt ggc gcg atc     1104
115 Met Thr Val Asn Gly Lys Thr Ile Gly Asp Asn Cys Arg Gly Ala Ile
116         355                 360                 365
118 atc gaa gac gag aag gtc atc cgc ccc tat gag cag ccg ctc aag gag     1152
119 Ile Glu Asp Glu Lys Val Ile Arg Pro Tyr Glu Gln Pro Leu Lys Glu
120     370                 375                 380
122 cgt gcc ggc ttc cgc gtt ctg cgc ggc aat ctg ttc tcc tcg gcc atc     1200
123 Arg Ala Gly Phe Arg Val Leu Arg Gly Asn Leu Phe Ser Ser Ala Ile
124 385                 390                 395                 400
```

```
126 atg aag aca agc gtg att tcg gaa gaa ttc cgc ggt cgt tac ctc tcc      1248
127 Met Lys Thr Ser Val Ile Ser Glu Glu Phe Arg Gly Arg Tyr Leu Ser
128             405                 410                 415
130 aac cct gat gat ccg gaa gcc ttc gaa ggc cgc gcc gtg gtg ttc gat      1296
131 Asn Pro Asp Asp Pro Glu Ala Phe Glu Gly Arg Ala Val Val Phe Asp
132             420                 425                 430
134 ggt ccg gag gat tac cat cat cgc atc gac gat ccg tcg ctt ggc atc      1344
135 Gly Pro Glu Asp Tyr His His Arg Ile Asp Asp Pro Ser Leu Gly Ile
136             435                 440                 445
138 gac gcc aac acc gtc ctg ttc atg cgc ggc gcc ggt ccg atc ggt tac      1392
139 Asp Ala Asn Thr Val Leu Phe Met Arg Gly Ala Gly Pro Ile Gly Tyr
140     450                 455                 460
142 ccg ggc gca gcg gaa gtg gtg aac atg cgc gcg ccg gat tac ctt ctg      1440
143 Pro Gly Ala Ala Glu Val Val Asn Met Arg Ala Pro Asp Tyr Leu Leu
144 465                 470                 475                 480
146 aag caa ggc gtc agt tcg ctg ccc tgc atc ggc gat ggc cgc cag tcc      1488
147 Lys Gln Gly Val Ser Ser Leu Pro Cys Ile Gly Asp Gly Arg Gln Ser
148             485                 490                 495
150 ggc acg tcg ggc agc cca tcc atc ctc aat gcc tcg ccg gaa gcg gcg      1536
151 Gly Thr Ser Gly Ser Pro Ser Ile Leu Asn Ala Ser Pro Glu Ala Ala
152             500                 505                 510
154 gcc ggc ggc ggt ctg tct att ctg cag acg ggt gac cgc gtc cgc atc      1584
155 Ala Gly Gly Gly Leu Ser Ile Leu Gln Thr Gly Asp Arg Val Arg Ile
156             515                 520                 525
158 gat gtg ggc cgc ggc aag gcc gat atc ctg ata tca ggt gaa gag ctc      1632
159 Asp Val Gly Arg Gly Lys Ala Asp Ile Leu Ile Ser Gly Glu Glu Leu
160     530                 535                 540
162 gcc aag cgt tac gag gcg ctg gca gct cag ggc ggt tat aag ttc ccc      1680
163 Ala Lys Arg Tyr Glu Ala Leu Ala Ala Gln Gly Gly Tyr Lys Phe Pro
164 545                 550                 555                 560
166 gac cac cag acg ccg tgg cag gaa atc cag cgc ggt atc gtc agc cag      1728
167 Asp His Gln Thr Pro Trp Gln Glu Ile Gln Arg Gly Ile Val Ser Gln
168             565                 570                 575
170 atg gaa acc ggc gcg gtt ctg gaa ccg gcc gta aag tat cag cgc atc      1776
171 Met Glu Thr Gly Ala Val Leu Glu Pro Ala Val Lys Tyr Gln Arg Ile
172             580                 585                 590
174 gcc cag acc aag ggc ctg ccg cgc gat aac cac tga                      1812
175 Ala Gln Thr Lys Gly Leu Pro Arg Asp Asn His
176             595                 600
179 <210> SEQ ID NO: 2
180 <211> LENGTH: 603
181 <212> TYPE: PRT
182 <213> ORGANISM: Agrobacterium tumefaciens
184 <400> SEQUENCE: 2
186 Met Thr Thr Ser Asp Asn Leu Pro Ala Thr Gln Gly Lys Leu Arg Ser
187 1               5                   10                  15
190 Arg Ala Trp Phe Asp Asn Pro Ala Asn Ala Asp Met Thr Ala Leu Tyr
191             20                  25                  30
194 Leu Glu Arg Tyr Met Asn Phe Gly Leu Ser Gln Ala Glu Leu Gln Ser
```

```
195           35                    40                    45
198 Asp Arg Pro Ile Ile Gly Ile Ala Gln Thr Gly Ser Asp Leu Ser Pro
199     50                    55                    60
202 Cys Asn Arg His His Leu Glu Leu Ala Asn Arg Leu Arg Glu Gly Ile
203 65                    70                    75                    80
206 Arg Glu Ala Gly Gly Ile Ala Ile Glu Phe Pro Val His Pro Ile Gln
207                   85                    90                    95
210 Glu Thr Gly Lys Arg Pro Thr Ala Gly Leu Asp Arg Asn Leu Ala Tyr
211                   100                   105                   110
214 Leu Gly Leu Val Glu Val Leu Tyr Gly Tyr Pro Leu Asp Gly Val Val
215         115                   120                   125
218 Leu Thr Ile Gly Cys Asp Lys Thr Thr Pro Ala Cys Leu Met Ala Ala
219     130                   135                   140
222 Ala Thr Val Asn Ile Pro Ala Ile Ala Leu Ser Val Gly Pro Met Leu
223 145                   150                   155                   160
226 Asn Gly Trp Phe Arg Gly Glu Arg Thr Gly Ser Gly Thr Ile Val Trp
227                   165                   170                   175
230 Lys Ala Arg Glu Leu Leu Ala Lys Gly Glu Ile Asp Tyr Gln Gly Phe
231             180                   185                   190
234 Val Lys Leu Val Ala Ser Ser Ala Pro Ser Thr Gly Tyr Cys Asn Thr
235         195                   200                   205
238 Met Gly Thr Ala Thr Thr Met Asn Ser Leu Ala Glu Ala Leu Gly Met
239     210                   215                   220
242 Gln Leu Pro Gly Ser Ala Ala Ile Pro Ala Pro Tyr Arg Asp Arg Gln
243 225                   230                   235                   240
246 Glu Val Ser Tyr Leu Thr Gly Leu Arg Ile Val Asp Met Val Arg Glu
247             245                   250                   255
250 Asp Leu Lys Pro Ser Asp Ile Met Thr Lys Asp Ala Phe Ile Asn Ala
251         260                   265                   270
254 Ile Arg Val Asn Ser Ala Ile Gly Gly Ser Thr Asn Ala Pro Ile His
255     275                   280                   285
258 Leu Asn Gly Leu Ala Arg His Val Gly Val Glu Leu Thr Val Asp Asp
259     290                   295                   300
262 Trp Gln Thr Tyr Gly Glu Asp Val Pro Leu Leu Val Asn Leu Gln Pro
263 305                   310                   315                   320
266 Ala Gly Glu Tyr Leu Gly Glu Asp Tyr Tyr His Ala Gly Gly Val Pro
267                   325                   330                   335
270 Ala Val Val Asn Gln Leu Met Thr Gln Gly Leu Ile Met Glu Asp Ala
271         340                   345                   350
274 Met Thr Val Asn Gly Lys Thr Ile Gly Asp Asn Cys Arg Gly Ala Ile
275     355                   360                   365
278 Ile Glu Asp Glu Lys Val Ile Arg Pro Tyr Glu Gln Pro Leu Lys Glu
279     370                   375                   380
282 Arg Ala Gly Phe Arg Val Leu Arg Gly Asn Leu Phe Ser Ser Ala Ile
283 385                   390                   395                   400
286 Met Lys Thr Ser Val Ile Ser Glu Glu Phe Arg Gly Arg Tyr Leu Ser
287             405                   410                   415
290 Asn Pro Asp Asp Pro Glu Ala Phe Glu Gly Arg Ala Val Val Phe Asp
291         420                   425                   430
```

```
294 Gly Pro Glu Asp Tyr His His Arg Ile Asp Asp Pro Ser Leu Gly Ile
295         435                 440                 445
298 Asp Ala Asn Thr Val Leu Phe Met Arg Gly Ala Gly Pro Ile Gly Tyr
299         450                 455                 460
302 Pro Gly Ala Ala Glu Val Val Asn Met Arg Ala Pro Asp Tyr Leu Leu
303 465 ·             470                 475                 480
306 Lys Gln Gly Val Ser Ser Leu Pro Cys Ile Gly Asp Gly Arg Gln Ser
307                 485                 490                 495
310 Gly Thr Ser Gly Ser Pro Ser Ile Leu Asn Ala Ser Pro Glu Ala Ala
311         500                 505                 510
314 Ala Gly Gly Gly Leu Ser Ile Leu Gln Thr Gly Asp Arg Val Arg Ile
315         515                 520                 525
318 Asp Val Gly Arg Gly Lys Ala Asp Ile Leu Ile Ser Gly Glu Glu Leu
319         530                 535                 540
322 Ala Lys Arg Tyr Glu Ala Leu Ala Ala Gln Gly Gly Tyr Lys Phe Pro
323 545                 550                 555                 560
326 Asp His Gln Thr Pro Trp Gln Glu Ile Gln Arg Gly Ile Val Ser Gln
327                 565                 570                 575
330 Met Glu Thr Gly Ala Val Leu Glu Pro Ala Val Lys Tyr Gln Arg Ile
331         580                 585                 590
334 Ala Gln Thr Lys Gly Leu Pro Arg Asp Asn His
335         595                 600
338 <210> SEQ ID NO: 3
339 <211> LENGTH: 1272
340 <212> TYPE: DNA
341 <213> ORGANISM: Agrobacterium tumefaciens
343 <220> FEATURE:
344 <221> NAME/KEY: CDS
345 <222> LOCATION: (1)..(1269)
346 <223> OTHER INFORMATION:
349 <400> 3
350 atg cag tct tct tca gct ctt cgg caa tca acc ggc gat cag tcg gaa      48
351 Met Gln Ser Ser Ser Ala Leu Arg Gln Ser Thr Gly Asp Gln Ser Glu
352 1               5                   10                  15
354 tac cat gcc cag tcg aat atg atc ggc tct agc ccg gcg gac ggt ttg      96
355 Tyr His Ala Gln Ser Asn Met Ile Gly Ser Ser Pro Ala Asp Gly Leu
356             20                  25                  30
358 ctc gca ttg ccg ctt ctg acc gtc gat ctt gcc gtc tat cgc ggt aat     144
359 Leu Ala Leu Pro Leu Leu Thr Val Asp Leu Ala Val Tyr Arg Gly Asn
360             35                  40                  45
362 cgg gat cgc ttt ctt gcg ctt gtc tcg gcc cat gga gcg aag gcg gct     192
363 Arg Asp Arg Phe Leu Ala Leu Val Ser Ala His Gly Ala Lys Ala Ala
364     50                  55                  60
366 cca cat gcc aag acg ccg atg tgc ccg gag atc gcg atc gat ctg att     240
367 Pro His Ala Lys Thr Pro Met Cys Pro Glu Ile Ala Ile Asp Leu Ile
368 65                  70                  75                  80
370 gaa gcc ggt gcc tgg ggc gcg acg gtc gcc gat ctc ttc cag gcg gaa     288
371 Glu Ala Gly Ala Trp Gly Ala Thr Val Ala Asp Leu Phe Gln Ala Glu
372             85                  90                  95
```

W--> 349 <400> 3

**VERIFICATION SUMMARY**                DATE: 02/22/2007
PATENT APPLICATION: **US/10/560,760**      TIME: 12:11:03

Input Set : **A:\A04002\G3111EP.ST25.txt**
Output Set: **N:\CRF4\02222007\J560760.raw**

```
L:10 M:270 C: Current Application Number differs, Replaced Current Application No
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:25 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:22
L:349 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:346
```